

SEQUENCE LISTING

<110> SUGIURA, Takeyuki

<120> TETRAHYDROFOLATE SYNTHETASE GENE

<130> Q94181

<140> US 10/573,969

<141> 2006-03-30

<150> PCT/JP2004/014812

<151> 2004-09-30

<150> JP 2003-341245

<151> 2003-09-30

<160> 7

<170> PatentIn version 3.1

<210> 1

<211> 2934

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(2934)

<220>

<221> sig_peptide

<222> (1)..(93)

<400> 1

atg ggc acg cgt ctg ccg ctc gtc ctg cgc cag ctc cgc cgc ccg ccc	48
Met Gly Thr Arg Leu Pro Leu Val Leu Arg Gln Leu Arg Arg Pro Pro	
1 5 10 15	

cag ccc ccg ggc cct ccg cgc cgc ctc cgt gtg ccc tgt cgc gct agc	96
Gln Pro Pro Gly Pro Pro Arg Arg Leu Arg Val Pro Cys Arg Ala Ser	
20 25 30	

agc ggc ggc ggc gga ggc ggc ggc ggt ggc cgg gag ggc ctg ctt gga	144
Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Arg Glu Gly Leu Leu Gly	
35 40 45	

cag cgg cgg ccg cag gat ggc cag gcc cgg agc agc tgc agc ccc ggc	192
Gln Arg Arg Pro Gln Asp Gly Gln Ala Arg Ser Ser Cys Ser Pro Gly	
50 55 60	

ggc cga acg ccc gcg gcg cgg gac tcc atc gtc aga gaa gtc att cag	240
Gly Arg Thr Pro Ala Ala Arg Asp Ser Ile Val Arg Glu Val Ile Gln	
65 70 75 80	

aat tca aaa gaa gtt cta agt tta ttg caa gaa aaa aac cct gcc ttc	288
Asn Ser Lys Glu Val Leu Ser Leu Leu Gln Glu Lys Asn Pro Ala Phe	
85 90 95	

aag ccg gtt ctt gca att atc cag gca ggt gac gac aac ttg atg cag	336
-----------------------------------------------------------------	-----

(1/10)

Lys	Pro	Val	Leu	Ala	Ile	Ile	Gln	Ala	Gly	Asp	Asp	Asn	Leu	Met	Gln		
			100					105					110				
gaa	atc	aac	cag	aat	ttg	gct	gag	gag	gct	ggg	ctg	aac	atc	act	cac	384	
Glu	Ile	Asn	Gln	Asn	Leu	Ala	Glu	Glu	Ala	Gly	Leu	Asn	Ile	Thr	His		
		115					120					125					
att	tgc	ctc	cct	cca	gat	agc	agt	gaa	gcc	gag	att	ata	gat	gaa	atc	432	
Ile	Cys	Leu	Pro	Pro	Asp	Ser	Ser	Glu	Ala	Glu	Ile	Ile	Asp	Glu	Ile		
	130					135					140						
tta	aag	atc	aat	gaa	gat	acc	aga	gta	cat	ggc	ctt	gcc	ctt	cag	atc	480	
Leu	Lys	Ile	Asn	Glu	Asp	Thr	Arg	Val	His	Gly	Leu	Ala	Leu	Gln	Ile		
145					150					155					160		
tct	gag	aac	ttg	ttt	agc	aac	aaa	gtc	ctc	aat	gcc	ttg	aaa	cca	gaa	528	
Ser	Glu	Asn	Leu	Phe	Ser	Asn	Lys	Val	Leu	Asn	Ala	Leu	Lys	Pro	Glu		
				165					170					175			
aaa	gat	gtg	gat	gga	gta	aca	gac	ata	aac	ctg	ggg	aag	ctg	gtg	cga	576	
Lys	Asp	Val	Asp	Gly	Val	Thr	Asp	Ile	Asn	Leu	Gly	Lys	Leu	Val	Arg		
			180					185					190				
ggg	gat	gcc	cat	gaa	tgt	ttt	gtt	tca	cct	gtt	gcc	aaa	gct	gta	att	624	
Gly	Asp	Ala	His	Glu	Cys	Phe	Val	Ser	Pro	Val	Ala	Lys	Ala	Val	Ile		
		195					200					205					
gaa	ctt	ctt	gaa	aaa	tca	ggg	gtc	aac	cta	gat	gga	aag	aag	att	ttg	672	
Glu	Leu	Leu	Glu	Lys	Ser	Gly	Val	Asn	Leu	Asp	Gly	Lys	Lys	Ile	Leu		
	210					215					220						
gta	gtg	ggg	gcc	cat	ggg	tct	ttg	gaa	gct	gct	cta	caa	tgc	ctg	ttc	720	
Val	Val	Gly	Ala	His	Gly	Ser	Leu	Glu	Ala	Ala	Leu	Gln	Cys	Leu	Phe		
225					230					235					240		
cag	aga	aaa	ggg	tcc	atg	aca	atg	agc	atc	cag	tgg	aaa	aca	cgc	cag	768	
Gln	Arg	Lys	Gly	Ser	Met	Thr	Met	Ser	Ile	Gln	Trp	Lys	Thr	Arg	Gln		
				245					250					255			
ctt	caa	agc	aag	ctt	cac	gag	gct	gac	att	gtg	gtc	cta	ggc	tca	cct	816	
Leu	Gln	Ser	Lys	Leu	His	Glu	Ala	Asp	Ile	Val	Val	Leu	Gly	Ser	Pro		
			260					265					270				
aag	cca	gaa	gag	att	ccc	ctt	act	tgg	ata	caa	cca	gga	act	act	gtt	864	
Lys	Pro	Glu	Glu	Ile	Pro	Leu	Thr	Trp	Ile	Gln	Pro	Gly	Thr	Thr	Val		
		275					280					285					
ctc	aac	tgc	tcc	cat	gac	ttc	ctg	tca	ggg	aag	gtt	ggg	tgt	ggc	tct	912	
Leu	Asn	Cys	Ser	His	Asp	Phe	Leu	Ser	Gly	Lys	Val	Gly	Cys	Gly	Ser		
	290					295					300						
cca	aga	ata	cat	ttt	ggg	gga	ctc	att	gag	gaa	gat	gat	gtg	att	ctc	960	
Pro	Arg	Ile	His	Phe	Gly	Gly	Leu	Ile	Glu	Glu	Asp	Asp	Val	Ile	Leu		
305					310					315					320		
ctt	gct	gca	gct	ctg	cga	att	cag	aac	atg	gtc	agt	agt	gga	agg	aga	1008	
Leu	Ala	Ala	Ala	Leu	Arg	Ile	Gln	Asn	Met	Val	Ser	Ser	Gly	Arg	Arg		
				325				330					335				
tgg	ctt	cgt	gaa	cag	cag	cac	agg	cgg	tgg	aga	ctt	cac	tgc	ttg	aaa	1056	

Trp	Leu	Arg	Glu	Gln	Gln	His	Arg	Arg	Trp	Arg	Leu	His	Cys	Leu	Lys	
			340					345					350			
ctt	cag	cct	ctc	tcc	cct	gtg	cca	agt	gac	att	gag	att	tca	aga	gga	1104
Leu	Gln	Pro	Leu	Ser	Pro	Val	Pro	Ser	Asp	Ile	Glu	Ile	Ser	Arg	Gly	
		355					360				365					
caa	act	cca	aaa	gct	gtg	gat	gtc	ctt	gcc	aag	gag	att	gga	ttg	ctt	1152
Gln	Thr	Pro	Lys	Ala	Val	Asp	Val	Leu	Ala	Lys	Glu	Ile	Gly	Leu	Leu	
	370					375					380					
gca	gat	gaa	att	gaa	atc	tat	ggc	aaa	agc	aaa	gcc	aaa	gta	cgt	ttg	1200
Ala	Asp	Glu	Ile	Glu	Ile	Tyr	Gly	Lys	Ser	Lys	Ala	Lys	Val	Arg	Leu	
385					390					395					400	
tcc	gtg	cta	gaa	agg	tta	aag	gat	caa	gca	gat	gga	aaa	tac	gtc	tta	1248
Ser	Val	Leu	Glu	Arg	Leu	Lys	Asp	Gln	Ala	Asp	Gly	Lys	Tyr	Val	Leu	
				405				410					415			
gtt	gct	ggg	atc	aca	ccc	acc	cct	ctt	gga	gaa	ggg	aag	agc	aca	gtc	1296
Val	Ala	Gly	Ile	Thr	Pro	Thr	Pro	Leu	Gly	Glu	Gly	Lys	Ser	Thr	Val	
		420						425					430			
acc	atc	ggg	ctt	gtg	cag	gct	ctg	acc	gca	cac	ctg	aat	gtc	aac	tcc	1344
Thr	Ile	Gly	Leu	Val	Gln	Ala	Leu	Thr	Ala	His	Leu	Asn	Val	Asn	Ser	
		435					440					445				
ttt	gcc	tgc	ttg	agg	cag	cct	tcc	caa	gga	ccg	acg	ttt	gga	gtg	aaa	1392
Phe	Ala	Cys	Leu	Arg	Gln	Pro	Ser	Gln	Gly	Pro	Thr	Phe	Gly	Val	Lys	
	450					455					460					
gga	gga	gcc	gcg	ggt	ggt	gga	tat	gcc	cag	gtc	atc	ccc	atg	gag	gag	1440
Gly	Gly	Ala	Ala	Gly	Gly	Gly	Tyr	Ala	Gln	Val	Ile	Pro	Met	Glu	Glu	
465					470				475						480	
ttc	aac	ctt	cac	ttg	act	gga	gac	atc	cac	gcc	atc	acc	gct	gcc	aat	1488
Phe	Asn	Leu	His	Leu	Thr	Gly	Asp	Ile	His	Ala	Ile	Thr	Ala	Ala	Asn	
				485					490					495		
aac	ttg	ctg	gct	gcc	gcc	atc	gac	acg	agg	att	ctt	cat	gaa	aac	acg	1536
Asn	Leu	Leu	Ala	Ala	Ala	Ile	Asp	Thr	Arg	Ile	Leu	His	Glu	Asn	Thr	
			500					505					510			
caa	aca	gat	aag	gct	ctg	tat	aat	cgg	ctg	gtt	cct	tta	gtg	aat	ggt	1584
Gln	Thr	Asp	Lys	Ala	Leu	Tyr	Asn	Arg	Leu	Val	Pro	Leu	Val	Asn	Gly	
		515					520					525				
gtc	aga	gaa	ttt	tca	gaa	att	cag	ctt	gct	cgg	cta	aaa	aaa	ctg	gga	1632
Val	Arg	Glu	Phe	Ser	Glu	Ile	Gln	Leu	Ala	Arg	Leu	Lys	Lys	Leu	Gly	
	530					535					540					
ata	aat	aag	act	gat	ccg	agc	aca	ctg	aca	gaa	gag	gaa	gtg	agt	aaa	1680
Ile	Asn	Lys	Thr	Asp	Pro	Ser	Thr	Leu	Thr	Glu	Glu	Glu	Val	Ser	Lys	
545					550					555					560	
ttt	gcc	cgt	ctc	gac	atc	gac	cca	tct	acc	atc	acg	tgg	cag	aga	gta	1728
Phe	Ala	Arg	Leu	Asp	Ile	Asp	Pro	Ser	Thr	Ile	Thr	Trp	Gln	Arg	Val	
				565				570					575			
ttg	gat	aca	aat	gac	cga	ttt	cta	cga	aaa	ata	acc	atc	ggg	cag	gga	1776

Leu	Asp	Thr	Asn	Asp	Arg	Phe	Leu	Arg	Lys	Ile	Thr	Ile	Gly	Gln	Gly	
			580					585					590			
aac	aca	gag	aag	ggc	cat	tac	cgg	cag	gcg	cag	ttt	gac	atc	gca	gtg	1824
Asn	Thr	Glu	Lys	Gly	His	Tyr	Arg	Gln	Ala	Gln	Phe	Asp	Ile	Ala	Val	
		595					600					605				
gcc	agc	gag	atc	atg	gcg	gtg	ctg	gcc	ctg	acg	gac	agc	ctc	gca	gac	1872
Ala	Ser	Glu	Ile	Met	Ala	Val	Leu	Ala	Leu	Thr	Asp	Ser	Leu	Ala	Asp	
	610					615					620					
atg	aag	gca	cgg	ctg	gga	agg	atg	gtg	gtg	gcc	agt	gac	aaa	agc	ggg	1920
Met	Lys	Ala	Arg	Leu	Gly	Arg	Met	Val	Val	Ala	Ser	Asp	Lys	Ser	Gly	
625					630					635					640	
cag	cct	gtg	aca	gca	gat	gat	ttg	ggg	gtg	aca	ggg	gct	ttg	aca	gtt	1968
Gln	Pro	Val	Thr	Ala	Asp	Asp	Leu	Gly	Val	Thr	Gly	Ala	Leu	Thr	Val	
				645					650					655		
ttg	atg	aaa	gat	gca	ata	aaa	cca	aac	ctg	atg	cag	acc	ctg	gaa	ggg	2016
Leu	Met	Lys	Asp	Ala	Ile	Lys	Pro	Asn	Leu	Met	Gln	Thr	Leu	Glu	Gly	
			660					665					670			
aca	cct	gtg	ttc	gtg	cat	gcg	ggc	cct	ttt	gct	aac	att	gct	cac	ggc	2064
Thr	Pro	Val	Phe	Val	His	Ala	Gly	Pro	Phe	Ala	Asn	Ile	Ala	His	Gly	
		675					680					685				
aac	tct	tca	gtg	ttg	gct	gat	aaa	att	gcc	ctg	aaa	ctg	gtt	ggg	gaa	2112
Asn	Ser	Ser	Val	Leu	Ala	Asp	Lys	Ile	Ala	Leu	Lys	Leu	Val	Gly	Glu	
	690					695					700					
gaa	gga	ttt	gta	gtg	acc	gaa	gct	ggc	ttt	ggg	gct	gac	atc	gga	atg	2160
Glu	Gly	Phe	Val	Val	Thr	Glu	Ala	Gly	Phe	Gly	Ala	Asp	Ile	Gly	Met	
705					710					715					720	
gag	aaa	ttc	ttc	aac	atc	aag	tgc	cga	gct	tcc	ggc	ttg	gtg	ccc	aac	2208
Glu	Lys	Phe	Phe	Asn	Ile	Lys	Cys	Arg	Ala	Ser	Gly	Leu	Val	Pro	Asn	
				725					730					735		
gtg	gtt	gtg	tta	gtg	gca	acg	gtg	cga	gct	ctg	aag	atg	cat	gga	ggc	2256
Val	Val	Val	Leu	Val	Ala	Thr	Val	Arg	Ala	Leu	Lys	Met	His	Gly	Gly	
			740					745					750			
ggg	cca	agt	gta	acg	gct	ggg	gtt	cct	ctt	aag	aaa	gaa	tat	aca	gag	2304
Gly	Pro	Ser	Val	Thr	Ala	Gly	Val	Pro	Leu	Lys	Lys	Glu	Tyr	Thr	Glu	
		755					760					765				
gag	aac	atc	cag	ctg	gtg	gca	gac	ggc	tgc	tgt	aac	ctc	cag	aag	caa	2352
Glu	Asn	Ile	Gln	Leu	Val	Ala	Asp	Gly	Cys	Cys	Asn	Leu	Gln	Lys	Gln	
	770					775					780					
att	cag	atc	act	cag	ctc	ttt	ggg	gtt	ccc	gtt	gtg	gtg	gct	ctg	aat	2400
Ile	Gln	Ile	Thr	Gln	Leu	Phe	Gly	Val	Pro	Val	Val	Val	Ala	Leu	Asn	
785					790					795					800	
gtc	ttc	aag	acc	gac	acc	cgc	gct	gag	att	gac	ttg	gtg	tgt	gag	ctt	2448
Val	Phe	Lys	Thr	Asp	Thr	Arg	Ala	Glu	Ile	Asp	Leu	Val	Cys	Glu	Leu	
				805					810					815		
gca	aag	cgg	gct	ggg	gcc	ttt	gat	gca	gtc	ccc	tgc	tat	cac	tgg	tcg	2496

35

40

45

Gln Arg Arg Pro Gln Asp Gly Gln Ala Arg Ser Ser Cys Ser Pro Gly
50 55 60

Gly Arg Thr Pro Ala Ala Arg Asp Ser Ile Val Arg Glu Val Ile Gln
65 70 75 80

Asn Ser Lys Glu Val Leu Ser Leu Leu Gln Glu Lys Asn Pro Ala Phe
85 90 95

Lys Pro Val Leu Ala Ile Ile Gln Ala Gly Asp Asp Asn Leu Met Gln
100 105 110

Glu Ile Asn Gln Asn Leu Ala Glu Glu Ala Gly Leu Asn Ile Thr His
115 120 125

Ile Cys Leu Pro Pro Asp Ser Ser Glu Ala Glu Ile Ile Asp Glu Ile
130 135 140

Leu Lys Ile Asn Glu Asp Thr Arg Val His Gly Leu Ala Leu Gln Ile
145 150 155 160

Ser Glu Asn Leu Phe Ser Asn Lys Val Leu Asn Ala Leu Lys Pro Glu
165 170 175

Lys Asp Val Asp Gly Val Thr Asp Ile Asn Leu Gly Lys Leu Val Arg
180 185 190

Gly Asp Ala His Glu Cys Phe Val Ser Pro Val Ala Lys Ala Val Ile
195 200 205

Glu Leu Leu Glu Lys Ser Gly Val Asn Leu Asp Gly Lys Lys Ile Leu
210 215 220

Val Val Gly Ala His Gly Ser Leu Glu Ala Ala Leu Gln Cys Leu Phe
225 230 235 240

Gln Arg Lys Gly Ser Met Thr Met Ser Ile Gln Trp Lys Thr Arg Gln
245 250 255

Leu Gln Ser Lys Leu His Glu Ala Asp Ile Val Val Leu Gly Ser Pro
260 265 270

Lys Pro Glu Glu Ile Pro Leu Thr Trp Ile Gln Pro Gly Thr Thr Val

(6/10)

275

280

285

Leu Asn Cys Ser His Asp Phe Leu Ser Gly Lys Val Gly Cys Gly Ser
 290 295 300

Pro Arg Ile His Phe Gly Gly Leu Ile Glu Glu Asp Asp Val Ile Leu
 305 310 315 320

Leu Ala Ala Ala Leu Arg Ile Gln Asn Met Val Ser Ser Gly Arg Arg
 325 330 335

Trp Leu Arg Glu Gln Gln His Arg Arg Trp Arg Leu His Cys Leu Lys
 340 345 350

Leu Gln Pro Leu Ser Pro Val Pro Ser Asp Ile Glu Ile Ser Arg Gly
 355 360 365

Gln Thr Pro Lys Ala Val Asp Val Leu Ala Lys Glu Ile Gly Leu Leu
 370 375 380

Ala Asp Glu Ile Glu Ile Tyr Gly Lys Ser Lys Ala Lys Val Arg Leu
 385 390 395 400

Ser Val Leu Glu Arg Leu Lys Asp Gln Ala Asp Gly Lys Tyr Val Leu
 405 410 415

Val Ala Gly Ile Thr Pro Thr Pro Leu Gly Glu Gly Lys Ser Thr Val
 420 425 430

Thr Ile Gly Leu Val Gln Ala Leu Thr Ala His Leu Asn Val Asn Ser
 435 440 445

Phe Ala Cys Leu Arg Gln Pro Ser Gln Gly Pro Thr Phe Gly Val Lys
 450 455 460

Gly Gly Ala Ala Gly Gly Gly Tyr Ala Gln Val Ile Pro Met Glu Glu
 465 470 475 480

Phe Asn Leu His Leu Thr Gly Asp Ile His Ala Ile Thr Ala Ala Asn
 485 490 495

Asn Leu Leu Ala Ala Ala Ile Asp Thr Arg Ile Leu His Glu Asn Thr
 500 505 510

Gln Thr Asp Lys Ala Leu Tyr Asn Arg Leu Val Pro Leu Val Asn Gly

(7/10)

515

520

525

Val Arg Glu Phe Ser Glu Ile Gln Leu Ala Arg Leu Lys Lys Leu Gly
 530 535 540

Ile Asn Lys Thr Asp Pro Ser Thr Leu Thr Glu Glu Glu Val Ser Lys
 545 550 555 560

Phe Ala Arg Leu Asp Ile Asp Pro Ser Thr Ile Thr Trp Gln Arg Val
 565 570 575

Leu Asp Thr Asn Asp Arg Phe Leu Arg Lys Ile Thr Ile Gly Gln Gly
 580 585 590

Asn Thr Glu Lys Gly His Tyr Arg Gln Ala Gln Phe Asp Ile Ala Val
 595 600 605

Ala Ser Glu Ile Met Ala Val Leu Ala Leu Thr Asp Ser Leu Ala Asp
 610 615 620

Met Lys Ala Arg Leu Gly Arg Met Val Val Ala Ser Asp Lys Ser Gly
 625 630 635 640

Gln Pro Val Thr Ala Asp Asp Leu Gly Val Thr Gly Ala Leu Thr Val
 645 650 655

Leu Met Lys Asp Ala Ile Lys Pro Asn Leu Met Gln Thr Leu Glu Gly
 660 665 670

Thr Pro Val Phe Val His Ala Gly Pro Phe Ala Asn Ile Ala His Gly
 675 680 685

Asn Ser Ser Val Leu Ala Asp Lys Ile Ala Leu Lys Leu Val Gly Glu
 690 695 700

Glu Gly Phe Val Val Thr Glu Ala Gly Phe Gly Ala Asp Ile Gly Met
 705 710 715 720

Glu Lys Phe Phe Asn Ile Lys Cys Arg Ala Ser Gly Leu Val Pro Asn
 725 730 735

Val Val Val Leu Val Ala Thr Val Arg Ala Leu Lys Met His Gly Gly
 740 745 750

Gly Pro Ser Val Thr Ala Gly Val Pro Leu Lys Lys Glu Tyr Thr Glu

(8/10)

755

760

765

Glu Asn Ile Gln Leu Val Ala Asp Gly Cys Cys Asn Leu Gln Lys Gln
 770 775 780

Ile Gln Ile Thr Gln Leu Phe Gly Val Pro Val Val Val Ala Leu Asn
 785 790 795 800

Val Phe Lys Thr Asp Thr Arg Ala Glu Ile Asp Leu Val Cys Glu Leu
 805 810 815

Ala Lys Arg Ala Gly Ala Phe Asp Ala Val Pro Cys Tyr His Trp Ser
 820 825 830

Val Gly Gly Lys Gly Ser Val Asp Leu Ala Arg Ala Val Arg Glu Ala
 835 840 845

Ala Ser Lys Arg Ser Arg Phe Gln Phe Leu Tyr Asp Val Gln Val Pro
 850 855 860

Ile Val Asp Lys Ile Arg Thr Ile Ala Gln Ala Val Tyr Gly Ala Lys
 865 870 875 880

Asp Ile Glu Leu Ser Pro Glu Ala Gln Ala Lys Ile Asp Arg Tyr Thr
 885 890 895

Gln Gln Gly Phe Gly Asn Leu Pro Ile Cys Met Ala Lys Thr His Leu
 900 905 910

Ser Leu Ser His Gln Pro Asp Lys Lys Gly Val Pro Arg Asp Phe Ile
 915 920 925

Leu Pro Ile Ser Asp Val Arg Ala Ser Ile Gly Ala Gly Phe Ile Tyr
 930 935 940

Pro Leu Val Gly Thr Met Ser Thr Met Pro Gly Leu Pro Thr Arg Pro
 945 950 955 960

Cys Phe Tyr Asp Ile Asp Leu Asp Thr Glu Thr Glu Gln Val Lys Gly
 965 970 975

Leu Phe

<210> 3

(9/10)

<211> 38
<212> DNA
<213> Homo sapiens

<400> 3
cgggatccgc catgggcacg cgtctgccgc tcgtcctg 38

<210> 4
<211> 38
<212> DNA
<213> Homo sapiens

<400> 4
ccgctcgagg aacaagcctt taacttggtc tgtttcgg 38

<210> 5
<211> 23
<212> DNA
<213> Homo sapiens

<400> 5
gctttggtgc tgacatcgga atg 23

<210> 6
<211> 23
<212> DNA
<213> Homo sapiens

<400> 6
cccggacgtc actgataggt aag 23

<210> 7
<211> 5
<212> PRT
<213> Homo sapiens

<400> 7

Ser Ser Gly Gly Gly
1 5